SI Appendix

Much theoretical work has been devoted to quantifying the conditions under which microscopic fluctuations have macroscopic effects [1]. The most useful results are often restricted to systems with a single degree of freedom or employ sophisticated tools such as Itô's calculus. In what follows, we aim to develop a convenient and simple scheme to assess the stability properties of a dynamical system subject to molecular noise described by the chemical Master equation. The method is an extension of the familiar linear stability analysis of nonlinear dynamical systems, although here the effective eigenvalues about the equilibrium points are adjusted to reflect the influence of the noise.

I. MATHEMATICAL METHODS

A very useful qualitative picture of the behavior of a system of nonlinear differential equations emerges from the linearized dynamics about the *fixed-point(s)* (also called the *steady-state(s)*) of the system, defined as the reactant concentrations at which the synthesis and degradation rates balance. The stability of the system near the fixed-points can be estimated by calculating the *eigenvalues* $\{\lambda_i\}$ of the resulting linearization, which are generally a set of complex numbers. If the real parts are all negative, we say the system is locally stable, meaning small perturbations away from the steady-state are automatically corrected.

Since genetic circuits, both natural and engineered, rely upon transfer of information through small numbers of molecules, significant fluctuation is simply one of the inherent operating conditions [2], resulting in noise that may give rise to behavior that is very different from the behavior predicted by deterministic models. Consequently, for cell-scale modeling we propose to modify the deterministic notion of stability by calculating the *effective* eigenvalues λ_i', which include the averaged influence of the intrinsic noise,

$$\lambda_i' = \lambda_i + \lambda_{corr}. \tag{1}$$

Here $\lambda_{corr} \propto V_{cell}^{-1}$ is inversely proportional to the cell volume V_{cell}. For notational convenience in the following, we introduce a parameter ω that is related to the cell volume by: $\omega^{-2} = V_{cell}$. Sometimes ω^{-2} is called the 'system size', expressing as it does the relationship between reactant concentration and molecule numbers [3, 4].

A. Stochastic stability equation

To calculate the stability of the macroscopic model $\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x})$ to small perturbations, the system is linearized about the equilibrium point: $\mathbf{x} = \mathbf{x}_s + \mathbf{x}_p$,

$$\frac{d}{dt}\mathbf{x}_p = \mathbf{J^{(0)}} \cdot \mathbf{x}_p. \tag{2}$$

(Here, and henceforth, we adopt the convention of writing all matrix variables in bold upper-case, and all vectors in bold lower-case.) The eigenvalues of the Jacobian $\mathbf{J}^{(0)} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\big|_{\mathbf{x}=\mathbf{x}_s}$ provide the decay rate of the exponential eigenmodes; if all the eigenvalues have negative real part, we say the system is *locally asymptotically stable*. We shall restrict ourselves to this notion of stability, although it does ignore *algebraically* growing modes which may be important in certain instances [5].

To accommodate fluctuations on top of the small perturbation $\mathbf{x}_p$, we set $\mathbf{x} = \mathbf{x}_s + \mathbf{x}_p + \omega\boldsymbol{\alpha}(t)$. The Jacobian

$$\mathbf{J} \equiv \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_s+\omega\boldsymbol{\alpha}},$$

will then be a (generally) nonlinear function of the fluctuations about the steady-state $\boldsymbol{\alpha}(t)$. (As a technical aside, we note that we are justified in replacing $\mathbf{x}$ by $\mathbf{x}_s + \mathbf{x}_p + \omega\boldsymbol{\alpha}(t)$ in both the right- *and* left-hand side of the deterministic model $\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x})$ since the fluctuations $\boldsymbol{\alpha}(t)$ have non-zero correlation time (as we show below) and zero mean, allowing us first to conclude that the time-derivative of $\boldsymbol{\alpha}(t)$ exists and further that the average of this derivative must vanish: $\langle\frac{d\boldsymbol{\alpha}}{dt}\rangle = \frac{d\langle\boldsymbol{\alpha}\rangle}{dt} = 0$). In the limit $\omega \to 0$, we can further linearize $\mathbf{J}$ with respect to ω,

$$\mathbf{J} \approx \mathbf{J}|_{\omega\to 0} + \omega \left.\frac{\partial \mathbf{J}}{\partial \omega}\right|_{\omega\to 0} \equiv \mathbf{J}^{(0)} + \omega\mathbf{J}^{(1)}(t).$$

The stability equation is then given by,

$$\frac{d}{dt}\mathbf{x}_p = [\mathbf{J^{(0)}} + \omega\mathbf{J}^{(1)}(t)] \cdot \mathbf{x}_p. \tag{3}$$

This is a linear stochastic differential equation with random coefficient matrix $\mathbf{J}^{(1)}(t)$ composed of a linear combination of the steady-state fluctuations $\boldsymbol{\alpha}(t)$ which have *non-zero* correlation time (see Eq. 12). We therefore need not appeal to any specialized calculi (*e.g.* Itô's calculus) for interpretation since the non-vanishing correlation time of the fluctuations ensures that $\mathbf{x}_p$ is a differentiable process and the equation falls under the purview of ordinary calculus [6].

Our present interest is in the *mean stability* of the equilibrium point. Taking the ensemble average of Eq. 3,

$$\frac{d}{dt}\langle\mathbf{x}_p\rangle = \mathbf{J}^{(0)} \cdot \langle\mathbf{x}_p\rangle + \omega\left\langle\mathbf{J}^{(1)}(t) \cdot \mathbf{x}_p\right\rangle.$$

The right-most term is the cross-correlation between the process $\mathbf{x}_p$ and the coefficient matrix $\mathbf{J}^{(1)}(t)$. Since the correlation time of $\mathbf{J}^{(1)}(t)$ is not small compared with the other time scales in the problem, it cannot be replaced by white noise, and an approximation scheme must be developed to find a closed evolution equation for $\langle \mathbf{x}_p \rangle$.

B. Bourret's mode-coupling approximation

By assumption, the number of molecules is large so the parameter ω is small, although not so small that intrinsic fluctuations can be ignored. To leading-order in ω, the trajectory $\mathbf{x}_p(t)$ is a random function of time since it is described by a differential equation with random coefficients. Derivation of the entire probability distribution of $\mathbf{x}_p(t)$ is usually impossible, and we must resort to methods of approximation. We shall adopt the closure scheme of Bourret [7–9] to arrive at a deterministic equation for the evolution of the averaged process $\langle \mathbf{x}_p(t) \rangle$ in terms of only the first and second moments of the fluctuations. In that approximation, provided $\mathbf{J}^{(0)} \gg \omega \mathbf{J}^{(1)}$, the dynamics of $\langle \mathbf{x}_p \rangle$ are governed by the convolution equation,

$$\frac{d}{dt}\langle \mathbf{x}_p(t) \rangle = \mathbf{J}_0 \langle \mathbf{x}_p(t) \rangle + \omega^2 \int_0^t \mathbf{J}_c(t-\tau) \langle \mathbf{x}_p(\tau) \rangle \, d\tau, \tag{4}$$

where $\mathbf{J}_c(t-\tau) = \left\langle \mathbf{J}^{(1)}(t) \, e^{\mathbf{J}^{(0)}(t-\tau)} \mathbf{J}^{(1)}(\tau) \right\rangle$ is the time autocorrelation matrix of the fluctuations and $e^{\mathbf{J}_0(t-\tau)}$ is the *matrix exponential* of $\mathbf{J}^{(0)}$. The equation can be solved formally by Laplace transform,

$$\langle \hat{\mathbf{x}}_\mathbf{p}(s) \rangle = \left[s\mathbf{I} - \mathbf{J}^{(0)} - \omega^2 \hat{\mathbf{J}}_c(s) \right]^{-1} \langle \mathbf{x}_p(0) \rangle,$$

where now $\hat{\mathbf{J}}_c(s) = \int_0^t \mathbf{J}_c(t) \, e^{-st} dt$. A necessary and sufficient condition for asymptotic stability of the averaged perturbation modes $\langle \mathbf{x}_p(t) \rangle$ is that the roots λ' of the resolvent,

$$\det \left[\lambda' \mathbf{I} - \mathbf{J}_0 - \omega^2 \hat{\mathbf{J}}_c(\lambda') \right] = 0, \tag{5}$$

all have negative real parts ($Re(\lambda') < 0$) [10]. Some insight into the behavior of the system can be gained by considering a perturbation expansion of the effective eigenvalues λ' in terms of the small parameter ω. We further diagonalize $\mathbf{J}^{(0)}$, $\mathrm{diag}[\lambda_i] = \mathbf{P}^{-1} \cdot \mathbf{J}^{(0)} \cdot \mathbf{P}$, and provided the eigenvalues are distinct, we can explicitly write λ_i' in terms of the unperturbed eigenvalues λ_i to $O(\omega^4)$ as,

$$\lambda_i' = \lambda_i + \omega^2 \left[\, \mathbf{P}^{-1} \cdot \hat{\mathbf{J}}_c(\lambda_i) \cdot \mathbf{P} \, \right]_{ii}, \tag{6}$$

where $[\,\cdot\,]_{ii}$ denotes the i^{th} diagonal entry of the matrix.

Notice the matrix product $\mathbf{J}_c(t-\tau)$ contains linear combinations of the correlation of the fluctuations $\langle \alpha_i(t) \alpha_j(\tau) \rangle$, and as such we must derive an expression for those moments.

C. Calculating the statistics of the steady-state fluctuations

The statistics of the fluctuations $\boldsymbol{\alpha}$ are fully determined by the solution of the chemical Master equation (defined below) that comes from treating each reaction event probabilistically. In that probabilistic formulation, our state at any time t is represented by the vector of molecule numbers $\mathbf{n} \in \mathbb{N}^d$; with n_i representing the number of molecules of a given species. Each reaction causes a transition from the initial state $\mathbf{n}$ to some new state $\mathbf{n}'$ reflecting the addition or removal of molecules by that reaction. The probability that the transition $\mathbf{n} \to \mathbf{n}'$ occurs is the product of the probability of being in state $\mathbf{n}$ at time t, $P(\mathbf{n}, t)$, and the transition probability of moving from $\mathbf{n} \to \mathbf{n}'$, denoted by $W_{\mathbf{n} \to \mathbf{n}'}$. We thus write the probability conservation as a balance of flux into and out of the state $\mathbf{n}$, which yields a discrete differential equation for $P(\mathbf{n}, t)$,

$$\frac{\partial P(\mathbf{n}, t)}{\partial t} = \sum_{\mathbf{n}'} W_{\mathbf{n}' \to \mathbf{n}} P(\mathbf{n}', t) - W_{\mathbf{n} \to \mathbf{n}'} P(\mathbf{n}, t). \tag{7}$$

The evolution equation for $P(\mathbf{n}, t)$ is called the Master equation [11]. It is rare that the Master equation can be solved exactly for $P(\mathbf{n}, t)$, and approximation schemes are required. One such scheme, the linear noise approximation [12], is versatile and will be described briefly (see also [3] and [13]). The approximation begins with the assumption that the molecule concentrations can be meaningfully separated into a component that evolves deterministically, which we shall denote $\mathbf{x}(t)$, and fluctuations $\boldsymbol{\alpha}(t)$ that account for the deviation of the stochastic model from the deterministic model. We introduce a scaling parameter ω, where $\omega^{-2} = V_{cell}$ is the volume of the cell and is an extensive measure of the number of molecules. We then make the ansatz that the fluctuations scale as the square-root of the number of molecules: $\omega^2 \, n_i = x_i + \omega \, \alpha_i$ [12, 14]. In that way, a perturbation expansion as the number of molecules gets large ($\omega \to 0$, with concentration held fixed), returns to zero'th order the macroscopic reaction rate equations,

$$\frac{d\mathbf{x}}{dt} = \boldsymbol{f}(\mathbf{x}). \tag{8}$$

The first-order equation, that comes at $O(\omega)$, characterizes the probability distribution for the fluctuations $\Pi(\boldsymbol{\alpha}, t)$ centered on the macroscopic trajectory $\mathbf{x}(t)$, and has the form of a *linear* Fokker-Planck equation,

$$\frac{\partial \Pi}{\partial t} = -\sum_{i,j} \Gamma_{ij} \partial_i (\alpha_j \Pi) + \frac{1}{2} \sum_{i,j} D_{ij} \partial_{ij} \Pi. \tag{9}$$

where ∂_i denotes $\partial/\partial\alpha_i$ and

$$\Gamma_{ij}(t) = \frac{\partial f_i}{\partial x_j} \qquad \boldsymbol{D} = \mathbf{S}\cdot\mathrm{diag}[\boldsymbol{\nu}]\cdot\mathbf{S}^T, \tag{10}$$

(see main text). The matrices $\boldsymbol{\Gamma}$ and $\boldsymbol{D}$ are independent of $\boldsymbol{\alpha}$, which appears only linearly in the drift term. As a consequence, the distribution $\Pi(\boldsymbol{\alpha}, t)$ will be Gaussian for all time. In particular, at equilibrium the fluctuations are distributed with density,

$$\Pi_s(\boldsymbol{\alpha}) = \left[(2\pi)^d \det\boldsymbol{\Xi}\right]^{\frac{1}{2}} \exp\left[-\frac{1}{2}\boldsymbol{\alpha}^T\cdot\boldsymbol{\Xi}^{-1}\cdot\boldsymbol{\alpha}\right],$$

and variance $\boldsymbol{\Xi} = \langle\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}^T\rangle$ determined by,

$$\boldsymbol{\Gamma}\cdot\boldsymbol{\Xi} + \boldsymbol{\Xi}\cdot\boldsymbol{\Gamma}^T + \mathbf{D} = 0. \tag{11}$$

Furthermore, the steady-state time correlation function is,

$$\left\langle\boldsymbol{\alpha}(t)\,\boldsymbol{\alpha}^T(t-\tau)\right\rangle = \exp[\boldsymbol{\Gamma}\tau]\cdot\boldsymbol{\Xi}. \tag{12}$$

Around the steady-state, the process is stationary, which means the correlation function depends upon time difference only. Also note that the characteristic correlation time $\tau_c = ||\boldsymbol{\Gamma}||^{-1}$ is related to the Jacobian $\boldsymbol{\Gamma}$ of the deterministic equations, and therefore *cannot* be divorced from the deterministic relaxation time. As a consequence, representing the fluctuations $\boldsymbol{\alpha}(t)$ as white noise ($\tau_c \to 0$) is *not* justified.

The great advantage of the linear noise approximation is that the autocorrelation function of the steady-state fluctuations can be calculated directly from the macroscopic reaction rates in an algorithmic fashion [3]. Furthermore, since $\boldsymbol{\Gamma}$ and $\mathbf{D}$ are derived from the known propensity and stoichiometry of the reactions, the statistics of $\boldsymbol{\alpha}$ are fully determined and are *not* tunable by some *ad hoc* prescription.

II. MEAN FIRST PASSAGE TIME

Bistability is a property exhibited by deterministic systems. In a stochastic context, bistability is sometimes assigned to an equilibrium probability distribution with two maxima, irrespective of their separation. A more practical criterion for bistability is that the two states are long-lived and that the mean escape time from one state to the other is longer than the natural timescales in the problem. For the single-variable autoactivator model, we are able to compute the escape time by an explicit (though approximate) expression (see [15] or p. 139 of [16] for details). Under fairly unrestrictive assumptions [17], the Master equation may be approximated by the nonlinear Fokker-Planck equation,

$$\frac{\partial P(a,t)}{\partial t} = -\frac{\partial}{\partial a}\Gamma(a)\,P(a,t) + \frac{1}{2}\frac{\partial^2}{\partial a^2}D(a)\,P(a,t),$$

where the functions Γ and D are the nonlinear analogues of the coefficient matrices $\boldsymbol{\Gamma}$ and $\mathbf{D}$ generated by the linear noise approximation shown in the previous section. For our autoactivator example, the coefficients are given by,

$$\Gamma(a) = \gamma\cdot g(a) - \delta\cdot a \qquad D(a) = \gamma\cdot b\cdot g(a) + \delta\cdot a.$$

The nonlinear Fokker-Planck equation has no general solution for systems of dimension greater than 1, and even the stationary solution is often impossible to calculate exactly for such systems [18]. In the reduced autoactivator model, we are fortunate to have a system with one independent variable, so we can write the stationary solution of the Fokker-Planck equation explicitly as,

$$P^s(a) = \frac{\mathcal{N}}{D(a)}\exp\left[2\int_0^a \frac{\Gamma(a')}{D(a')}da'\right],$$

where $\mathcal{N}$ is the constant of normalization (see p. 124 of [16]). Furthermore, we can explicitly write the first passage time τ from the *HIGH* state to the *LOW* state or vice-versa.

$$\tau_{HI\to LO} = 2\int_{a_{mid}}^{a^*_{HI}} \frac{1}{\psi(x)}\int_x^\infty \frac{\psi(y)}{D(y)}dydx$$

$$\tau_{LO\to HI} = 2\int_{a^\star_{LO}}^{a_{mid}} \frac{1}{\psi(x)}\int_0^x \frac{\psi(y)}{D(y)}dydx,$$

where a_{mid} is the unstable equilibrium point separating the *HIGH* and *LOW* states $a^\star_{HI}$ and $a^\star_{LO}$, respectively. The function $\psi(x)$ is given by,

$$\psi(x) = \exp\left[2\int_0^x \frac{\Gamma(x')}{D(x')}dx'\right],$$

(see p. 139 of [16] for additional details).

In the main text, we discuss $\min[\tau_{LO\to HI}, \tau_{HI\to LO}]$ along the stability curves predicted by the effective eigenvalues. For $\Delta_b = 0.1$, $\min[\tau_{LO\to HI}, \tau_{HI\to LO}] = 8 \pm 4$, where time has been scaled to protein lifetime (δ^{-1}). For $\Delta_b = 0.2$ and $\Delta_b = 0.3$, $\min[\tau_{LO\to HI}, \tau_{HI\to LO}] = 5.6 \pm 1.4$ and 5.9 ± 0.3, respectively.

III. DETAILS OF GENETIC CIRCUIT EXAMPLES

A. The autoactivator

We describe the transcription of the activator mRNA, m_a and the translation of activator protein A as two differential equations using the activation function g to describe the time-averaged state of the promoter,

$$\frac{dm_a}{dt} = \gamma_m\cdot g(A) - \delta_m\, m_a, \frac{dA}{dt} = \gamma_p\, m_a - \delta_p\, A. \tag{13}$$

Here γ_m is the transcription rate, γ_p is the translation rate, δ_m and δ_p are the rates of mRNA degradation and protein degradation, respectively. We make the assumption that the mRNA turnover is much faster than the timescale of protein degradation (*i.e.* $\delta_m \gg \delta_p$). In that way, we justify setting the mRNA concentration to its equilibrium level,

$$m^\star(A) = \frac{\gamma_m}{\delta_m}\, g(A), \tag{14}$$

reducing the model to a single equation,

$$\frac{dA}{dt} = \frac{\gamma_m \cdot \gamma_p}{\delta_m} \cdot g(A) - \delta_p\, A, \tag{15}$$

at the expense of lumping transcription and translation together. Re-writing the constants $\gamma = \frac{\gamma_m \cdot \gamma_p}{\delta_m}$ and $\delta_p = \delta$, we are left with the evolution equation as written in the main text,

$$\frac{dA}{dt} = \gamma \cdot g(A) - \delta \cdot A, \tag{16}$$

where γ is the fully activated rate of protein synthesis and δ is the rate of protein degradation.

1. *Transcriptional activation*

The lumping together of transcription and translation comes at the expense of obscuring translational amplification of the mRNA. The translational burst size is approximately equal to the averaged number of protein molecules synthesized during the lifetime of the mRNA, $b = \frac{\gamma_p}{\delta_m}$ [19, 20], so we see the production term in the macroscopic equation is actually ($b\times$ transcription rate),

$$\frac{dA}{dt} = b \times \gamma_m \cdot g(A) - \delta \cdot A. \tag{17}$$

In the deterministic model, the distinction between reaction rate and reaction stoichiometry is immaterial, but that is no longer true when we calculate the intrinsic fluctuations. Writing the production and degradation stoichiometry explicitly as in the main text,

$$\begin{array}{rl} \text{bursty synthesis:} & A \xrightarrow{\nu_1} A + b; \;\; \nu_1 = \frac{\gamma}{b} \cdot g\,(A), \\ \text{linear degradation:} & A \xrightarrow{\nu_2} A - 1; \quad \nu_2 = \delta \cdot A, \end{array} \tag{18}$$

leading to the propensity vector $\boldsymbol{\nu} = [\frac{\gamma}{b} \cdot g(A), \delta \cdot A]$ and stoichiometry matrix $\mathbf{S} = [b, -1]$. We can easily calculate the coefficient matrices $\boldsymbol{\Gamma}$ and $\mathbf{D}$,

$$\boldsymbol{\Gamma} = [\gamma \cdot g'(A) - \delta] \qquad \mathbf{D} = [b \cdot \gamma \cdot g(A) + \delta \cdot A]. \tag{19}$$

It is a simple task to then determine the steady-state correlations of the fluctuations,

$$\boldsymbol{\Xi} = -\frac{1}{2}\frac{\mathbf{D}}{\boldsymbol{\Gamma}} = -\frac{1}{2}\frac{[b \cdot \gamma \cdot g(A^\star) + \delta \cdot A^\star]}{[\gamma \cdot g'(A^\star) - \delta]}, \tag{20}$$

which is positive since the deterministic eigenvalue $\lambda = [\gamma\, g'(A^\star) - \delta] < 0$ in the stable regime where the analysis is carried out. We write the fractional deviation η of the steady-state fluctuations in A as,

$$\eta = \frac{\sqrt{\langle A^2 \rangle}}{A^\star} = \sqrt{\frac{(b+1)}{2\,[1 - A_0 g'\,(A^\star)]}}\sqrt{\frac{1}{A_0 \cdot V_{cell} \cdot g\,(A^\star)}},$$

where $A^\star$ is the steady-state activator concentration and $A_0 = \frac{\gamma}{\delta}$ is the fully-activated protein concentration and $\omega^{-2} = V_{cell}$ is the cell volume. Provided the *HIGH* and *LOW* equilibrium points are well-separated ($g'\,(A^\star) \approx 0$), we can write,

$$\eta_{LO} = \sqrt{\frac{(b+1)}{2}}\sqrt{\frac{f}{A_0 \cdot V_{cell}}} = \eta_{HI}\sqrt{f}, \tag{21}$$

where f is the *fold activation*. Not surprisingly, the relative fluctuations around the *LOW* state are large since in that state, the molecule numbers are small. More importantly for the present discussion, we see that the magnitude of the relative fluctuations depends directly upon the burstiness b. To determine the effect of the burstiness upon the averaged stability, we calculate the stability matrices $\mathbf{J}^{(0)}$ and $\mathbf{J}^{(1)}$ (where time has been scaled with respect to the protein lifetime: $t \to t \cdot \delta^{-1}$),

$$\mathbf{J}^{(0)} = [A_0\; g'_A(a) - 1] \qquad \omega\mathbf{J}^{(1)} = [A_0\; g''_A(a)]\omega\; \boldsymbol{\alpha}(t),$$

from which the Laplace transform of the autocorrelation function $\hat{\mathbf{J}}_c(s)$ is derived,

$$\omega^2\hat{\mathbf{J}}_c\,(s) = \omega^2\,[A_0 g'']^2 \int_0^\infty \langle \alpha\,(t)\,\alpha\,(0)\rangle\, e^{[A_0 g' - 1]t} e^{-st} dt.$$

Referring to Eq. 12, the steady-state fluctuations have exponential time-autocorrelation function so that the integrand becomes,

$$\omega^2\hat{\mathbf{J}}_c\,(s) = -\omega^2\,[A_0 g'']^2\,\frac{(b+1)}{2}\frac{A_0 g}{[A_0 g' - 1]} \times \int_0^\infty e^{[A_0 g' - 1]t} e^{[A_0 g' - 1]t} e^{-st} dt. \tag{22}$$

Evaluating the integral,

$$\omega^2\hat{\mathbf{J}}_c\,(s) = -\frac{(b+1)}{2}\frac{\omega^2}{K_A}\frac{A_0^2 g\,[A_0 g'']^2}{[A_0 g' - 1]}\frac{K_A}{A_0}\frac{1}{s - 2\,[A_0 g' - 1]}. \tag{23}$$

From the stability matrices, we are able to calculate the approximation of the effective eigenvalue λ' from Eq. 6,

$$\lambda' = [A_0 g' - 1] + \frac{\omega^2}{K_A}\frac{(b+1)}{2}\frac{K_A}{A_0}\frac{A_0^4\,[g'']^2\,g}{[A_0 g' - 1]^2}, \tag{24}$$

where we identify $\omega^{-2} = V_{cell}$ as the volume of the cell. Collecting the constants into groups, we write the the effective eigenvalue $\lambda'(A^\star)$ as,

$$\lambda' = \lambda + \frac{1}{V_{cell}}\lambda_{corr} = \lambda\left\{1 - \Delta_b \cdot h\left(\frac{A_0}{K_A}, g\left(A^\star\right)\right)\right\}, \tag{25}$$

where $\Delta_b = \frac{(b+1)}{2}\frac{1}{K_A \cdot V_{cell}}$ is the discrete change in reactant molecule numbers, scaled with respect to the number of activators required to initiate activation ($K_A \cdot V_{cell}$), representing the relative change in protein numbers incurred by the stochastic reaction events. (In a sense, K_A represents the characteristic concentration of the activator: for activator concentrations far less than K_A, there is no activation and for concentrations far above K_A, the promoter is fully activated.) The second term in Eq. 25, $h\left(\frac{A_0}{K_A}, g\left(A^\star\right)\right) = \frac{K_A}{A_0}\frac{A_0^4\left(g''\right)^2 g}{|\lambda|^3}$ contains the details of the regulatory mechanism [21] and depends strongly upon the stability of the deterministic system through λ. It is the interplay between the fluctuations (through Δ_b) and the macroscopic stability of the steady-state (through h) that ultimately decides the averaged stability of the stochastic system.

2. *Accuracy of ESA*

To compute the accuracy of the effective stability approximation as a function of the molecule numbers for the translational autoactivator model, the corrected eigenvalue λ' computed above (Eq. 25) is compared to the short-time Lyapunov exponent of the ensemble-averaged perturbation modes computed by stochastic simulation [22].

For a system slightly perturbed from the steady-state x_s, the short-time Lyapunov exponent $\langle\lambda\rangle$ is defined as,

$$\lim_{t\to 0} \ln\left|\langle x_p\left(t\right)\rangle - x_s\right| = \text{const.} + \langle\lambda\rangle \cdot t.$$

A numerical calculation of $\langle\lambda\rangle$ is obtained by taking the ensemble average (over an ensemble of 10^5 members) of $x_p(t)$ determined by stochastic simulation. The slope of the natural-log difference between the numerically generated perturbation mode and the steady state, $\ln|\langle x_p(t)\rangle - x_s|$, is fit by linear regression over a time span corresponding to the protein lifetime (*i.e.* $\delta^{-1} = 30$ minutes). To compare the stochastic simulation with the ESA, we focus upon three points in the parameter space of the autoactivator (Figure 1a, filled circles) – one point well inside the bistable regime ($\frac{A_0}{K_A} = 2.5, f = 80$; red), one near the boundary predicted by the ESA ($\frac{A_0}{K_A} = 3.5, f = 80$; green), and one well inside the monostable regime ($\frac{A_0}{K_A} = 5, f = 80$; blue). Figure 1b compares the resulting Lyapunov exponent $\langle\lambda\rangle$ (dashed lines) with the ESA prediction λ' (solid lines), where the line colors correspond to the colors of the filled circles

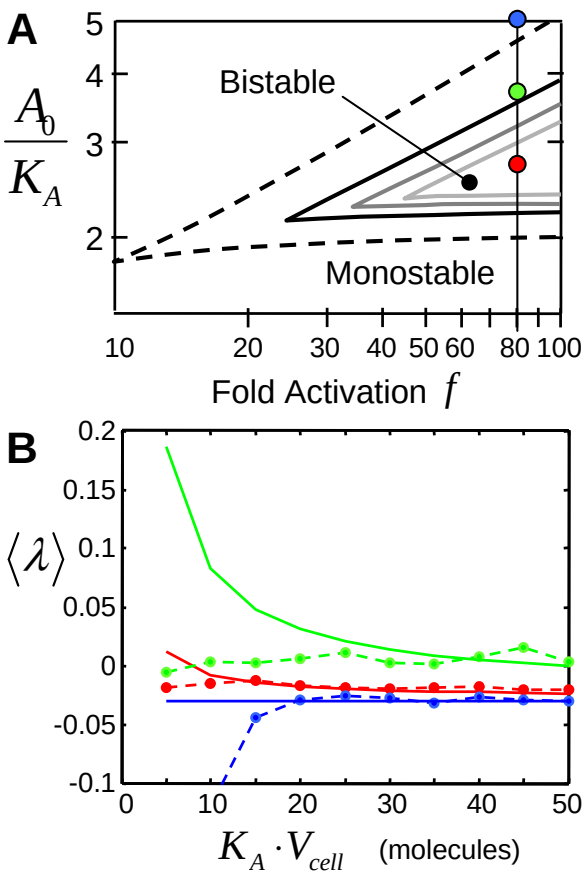

FIG. 1: Accuracy of the effective stability approximation (ESA) as a function of the number of molecules. (A) Focusing upon three points in the parameter space of the autoactivator model (see Figure 2a in the main text), it is possible to compare the ESA with the results of numerical simulation. (B) The short-time Lyapunov exponent of an ensemble average of the perturbation modes about the LOW state (dashed lines) approach those values of λ' predicted according to Eq. 25 (solid lines) for systems with increasing values of $K_A \cdot V_{cell}$, which specifies the order of molecule numbers to turn on/off the gene. Here, the burstiness of protein synthesis is held constant at $b = 9$, and each data point is computed from a sample of 10^5 trajectories – colors of the curves correspond to the filled circles in panel A.

in Figure 1a. Here, the burstiness in protein synthesis is held constant at $b = 9$, and the characteristic number of molecules in the system, $K_A \cdot V_{cell}$, is increased from 5 to 50. (In the main text, $K_A \cdot V_{cell} = 25$ so that a burstiness of $b = 9$ gives a discreteness parameter of $\Delta_{b_A} = \frac{(b+1)}{2}\frac{1}{K_A \cdot V_{cell}} = 0.2$.) As the number of molecules in the system is increased, the ESA and the numerical simulation results converge. The figure shows the effective stability of the transcriptional autoactivator model is well-characterized by the ESA for systems with $K_A \cdot V_{cell} \gtrsim 20$.

3. Translational activation

To model the translational activity, we redefine the transcription rate to be constant $\frac{\gamma}{b}$, where b is the maximum burst size at full activation, and allow the activator to control the translation rate through the *stoichiometery*. We write the synthesis and degradation reactions – in analogy with Eq. 18 above – as,

$$\begin{array}{cc} A \xrightarrow{\nu_1} A + b \cdot g\left(A\right); & \nu_1 = \frac{\gamma}{b}, \\ A \xrightarrow{\nu_2} A - 1; & \nu_2 = \delta \cdot A, \end{array} \tag{26}$$

where the translational activation affects the stoichiometry through the synthesis step-size $b \cdot g(A)$. Notice that the deterministic equation $\frac{dA}{dt} = \mathbf{S} \cdot \boldsymbol{\nu} = A_0\, g(A) - A$ is *identical* to the deterministic equation for the transcriptional autoactivator in the previous section. Nonetheless, the change in synthesis stoichiometry from $b \mapsto b \cdot g\left(A\right)$ has a noticeable effect on the resulting stability. As above, we calculate the effective eigenvalue,

$$\lambda' = \lambda \left\{ 1 - \frac{\left(b \cdot g\left(A^\star\right) + 1\right)}{2} \frac{1}{V_{cell} \cdot K_A} \cdot h\left(\frac{A_0}{K_A}, g\left(A^\star\right)\right) \right\},$$

where $h(\ \cdot\ )$ is as in Eq. 25. The difference from the transcriptional case is that the burst-size itself is attenuated in the *LOW* state, and the discreteness parameter approaches the minimal value $\Delta_b \to 1/(2V_{cell} \cdot K_A)$, thereby increasing the residence time in the *LOW* state as compared with transcriptional activation.

B. Genetic oscillator

The parameters of Vilar *et al.* [23] correspond to the reduced model parameters:

$$\begin{array}{c} \gamma_A = 25\ nM\ h^{-1}, K_A = 0.5\ nM, f_A = 10, \\ \gamma_R = 5\ nM\ h^{-1}, K_R = 1\ nM, f_R^{-1} = 0, \\ \kappa_C = 2 \times 10^2\ nM^{-1}\ h^{-1}, \text{and } \delta_A = 1\ h^{-1}, \end{array} \tag{27}$$

where, for simplicity, we make the approximation that 1 molecule / $1\mu m^3 \approx 1\ nM$ and set $V_{cell} = 100\mu m^3$. Furthermore, the mRNA degradation and translation rates in the original model give an activator burst size of $b_A = 5$ and a repressor burst size of $b_R = 10$.

1. Details of the stochastic model

The reduced model (Eq. 6 in the main text) is composed of six elementary reactions:

$$\begin{array}{cc} A \to A + b_A & \nu_1 = \frac{\gamma_A}{b_A} \cdot g\left(\frac{A}{K_A}, f_A\right) \\ A \to A - 1 & \nu_2 = \delta_A \cdot A \\ (A, R, C) \to (A - 1, R - 1, C + 1) & \nu_3 = \kappa_C \cdot A \cdot R \\ R \to R + b_R & \nu_4 = \frac{\gamma_R}{b_R} \cdot g\left(\frac{A}{K_R}, f_R\right) \\ R \to R - 1 & \nu_5 = \delta_R \cdot R \\ (R, C) \to (R + 1, C - 1) & \nu_6 = \delta_A \cdot C \end{array}$$

The stoichiometry matrix $\mathbf{S}$ and the propensity vector $\boldsymbol{\nu}$ are then written as,

$$\mathbf{S} = \begin{bmatrix} b_A & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & b_R & -1 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix}, \tag{28}$$

$$\boldsymbol{\nu} = \begin{bmatrix} \frac{\gamma_A}{b_A} \cdot g\left(\frac{A}{K_A}, f_A\right) \\ \delta_A \cdot A \\ \kappa_C \cdot A \cdot R \\ \frac{\gamma_R}{b_R} \cdot g\left(\frac{A}{K_R}, f_R\right) \\ \delta_R \cdot R \\ \delta_A \cdot C \end{bmatrix}.$$

Identification of dimensionless parameters in the deterministic model comes from considering the rate equations,

$$\frac{d}{dt}\begin{bmatrix} A \\ R \\ C \end{bmatrix} = \mathbf{S} \cdot \boldsymbol{\nu} = \tag{29}$$

$$\begin{bmatrix} \gamma_A \cdot g\left(\frac{A}{K_A}, f_A\right) - \delta_A \cdot A - \kappa_C \cdot A \cdot R \\ \gamma_R \cdot g\left(\frac{A}{K_R}, f_R\right) - \delta_R \cdot R - \kappa_C \cdot A \cdot R + \delta_A \cdot C \\ \kappa_C \cdot A \cdot R - \delta_A \cdot C \end{bmatrix}.$$

In what follows, it will be convenient to call $\gamma = \frac{\gamma_R}{\gamma_A}$ and $A_0 = \frac{\gamma_A}{\delta_A}$. Scaling the concentrations with respect to the characteristic concentration A_0 (*i.e.* $A = A' \cdot A_0$, *etc.*) and time with respect to the activator lifetime, $t = t' \cdot \delta_A$, the rate equations become,

$$\frac{d}{dt'}\begin{bmatrix} A' \\ R' \\ C' \end{bmatrix} = \tag{30}$$

$$\begin{bmatrix} g\left(A' \frac{A_0}{K_A}, f_A\right) - A' - \left[\frac{\kappa_C \cdot A_0}{\delta_A}\right] \cdot A' \cdot R' \\ \gamma \cdot g\left(A' \frac{A_0}{K_R}, f_R\right) - \left[\frac{\delta_R}{\delta_A}\right] \cdot R' - \left[\frac{\kappa_C \cdot A_0}{\delta_A}\right] \cdot A' \cdot R' + C' \\ \left[\frac{\kappa_C \cdot A_0}{\delta_A}\right] \cdot A' \cdot R' - C' \end{bmatrix}.$$

The two additional dimensionless constants are the scaled rate of dimerization $\kappa = \frac{\kappa_C \cdot A_0}{\delta_A}$ and the ratio of the repressor and activator degradation rates $\epsilon = \frac{\delta_R}{\delta_A}$. Henceforth, the primes denoting the dimensionless quantities will be dropped.

Since the variance in the fluctuations is found from the auxiliary matrices $\boldsymbol{\Gamma}$ and $\mathbf{D}$ (see Eq. 10), and $\boldsymbol{\Gamma}$ is the Jacobian of the deterministic system, the dimensionless stochastic parameters are most easily found by considering $\mathbf{D} = \mathbf{S} \cdot \mathrm{diag}\left[\boldsymbol{\nu}\right] \cdot \mathbf{S}^T$,

$$\mathbf{D} = \begin{bmatrix} b_A \cdot \gamma_A \cdot g_A + \delta_A \cdot A + \gamma_C \cdot A \cdot C & \gamma_C \cdot A \cdot C & -\gamma_C \cdot A \cdot C \\ \gamma_C \cdot A \cdot C & b_R \cdot \gamma_R \cdot g_R + \delta_R \cdot R + \gamma_C \cdot A \cdot C + \delta_A \cdot C & -\gamma_C \cdot A \cdot C - \delta_A \cdot C \\ -\gamma_C \cdot A \cdot C & -\gamma_C \cdot A \cdot C - \delta_A \cdot C & \gamma_C \cdot A \cdot C + \delta_A \cdot C \end{bmatrix},$$

where $g_i \equiv g\left(\frac{A}{K_i}, f_i\right)$. As above, we scale the concentrations with respect to A_0 and divide through by δ_A. Evaluating $\mathbf{D}$ at the steady-state $(A^\star, R^\star, C^\star)$, where $\frac{dA}{dt} = \frac{dR}{dt} = \frac{dC}{dt} = 0$, provides the additional simplifications derived from the rate equations above, written in dimensionless form,

$$\begin{aligned} g_A &= A^\star + \kappa \cdot A^\star \cdot R^\star, \\ \gamma \cdot g_R + C^\star &= \epsilon \cdot R^\star + \kappa \cdot A^\star \cdot R^\star, \\ C^\star &= \kappa \cdot A^\star \cdot R^\star. \end{aligned} \tag{31}$$

Hence, the matrix $\mathbf{D}$ is written in terms of reactant *numbers* as,

$$\frac{\mathbf{D}}{\gamma \cdot A_0} = \begin{bmatrix} 2\left[\frac{(b_A+1)}{2}\right]\frac{g_A}{\gamma} & C^\star & -C^\star \\ C^\star & 2\left[\frac{(b_R+1)}{2}\right]g_R + 2C^\star & -2C^\star \\ -C^\star & -2C^\star & 2C^\star \end{bmatrix}. \tag{32}$$

Comparing each diagonal element with the characteristic mean reactant number of that species ($N_A \sim K_A\ V_{cell}$, $N_R \sim K_R\ V_{cell}$), and ignoring parameters coming from the deterministic model ($g_A, g_R,$ and γ), we have three additional constants - the discreteness in the activator number $\Delta_{b_A} = \frac{(b_A+1)}{2}\frac{1}{K_A \cdot V_{cell}}$, the discreteness in the repressor number $\Delta_{b_R} = \frac{(b_R+1)}{2}\frac{1}{K_R \cdot V_{cell}}$ and the extent of dimerization $\frac{C^\star}{K_R \cdot V_{cell}}$. In the main text, we focus upon the effect of varying the deterministic parameter ϵ and the stochastic parameter Δ_{b_A}.

IV. ALGORITHMIC IMPLEMENTATION OF THE THE EFFECTIVE STABILITY APPROXIMATION

The corrections to the deterministic eigenvalues are computed by solving the resolvent equation for the the effective eigenvalues λ',

$$\det[\lambda' \cdot \mathbf{I} - \mathbf{J}^{(0)} - \frac{1}{V_{cell}}\hat{\mathbf{J}}_c(\lambda')], \tag{33}$$

(Eq. 12 in the main text). In this section, we provide a step-by-step algorithm to form the matrices $\mathbf{J}^{(0)}$ and $\hat{\mathbf{J}}_c(\lambda')$ from the deterministic reaction rates. In the following, the deterministic state vector is denoted by $\mathbf{x}$ and $\boldsymbol{\alpha}$ denotes the fluctuations in each of the components of $\mathbf{x}$ (see Section I-C above). The first three steps of the algorithm come from the paper by Elf and Ehrenberg [3].

1. Write the various reactions in terms of their *propensity* and *stoichiometry*. The deterministic reaction rates are defined by the product $\mathbf{S} \cdot \boldsymbol{\nu}$ (see Eqs. 18 and 28 above).

2. From $\mathbf{S}$ and $\boldsymbol{\nu}$, construct the matrices $\boldsymbol{\Gamma}$ and $\mathbf{D}$,

$$\boldsymbol{\Gamma}_{ij}(\mathbf{x}) = \frac{\partial[\mathbf{S} \cdot \boldsymbol{\nu}]_i}{\partial x_j} \qquad \mathbf{D}(\mathbf{x}) = \mathbf{S} \cdot \mathrm{diag}[\boldsymbol{\nu}] \cdot \mathbf{S}^T. \tag{34}$$

3. Compute the steady-state covariance in the fluctuations $\boldsymbol{\alpha}$ by solving the fluctuation-dissipation relation for each of the entries in the symmetric covariance matrix $\boldsymbol{\Xi}$ (where $\Xi_{ij} = \Xi_{ji} = \langle \alpha_i\ \alpha_j \rangle$),

$$\boldsymbol{\Gamma}(\mathbf{x}_s) \cdot \boldsymbol{\Xi} + \boldsymbol{\Xi} \cdot \boldsymbol{\Gamma}^T(\mathbf{x}_s) + \mathbf{D}(\mathbf{x}_s) = \mathbf{0}. \tag{35}$$

The steady-states $\mathbf{x}_s$ are calculated from the deterministic reaction rates by solving the algebraic equations $([\mathbf{S} \cdot \boldsymbol{\nu}]_{\mathbf{x}=\mathbf{x}_s}) = \mathbf{0}$.

Evaluated at the steady-state, the fluctuation-dissipation relation is simply a $\frac{1}{2}d(d+1)$ system of linear equations that determine the symmetric entries of $\boldsymbol{\Xi}$ (where d is the dimension of the system). For more details regarding the general solution of the fluctuation-dissipation relation, see [24].

4. Compute the matrices $\mathbf{J}^{(0)}$ and $\mathbf{J}^{(1)}(t)$,

$$\mathbf{J}^{(0)} = \boldsymbol{\Gamma}(\mathbf{x}_s) \qquad \mathbf{J}^{(1)}(t) = \frac{\partial \boldsymbol{\Gamma}(\mathbf{x}_s + \omega\ \boldsymbol{\alpha}(t))}{\partial \omega}|_{\omega=0}. \tag{36}$$

5. Calculate the matrix $\mathbf{J}_c(t)$,

$$\mathbf{J}_c(t) = \langle \mathbf{J}^{(1)}(t) \cdot \exp[\mathbf{J}^{(0)}\ t] \cdot \mathbf{J}^{(1)}(0) \rangle, \tag{37}$$

where $\exp[\mathbf{J}^{(0)}t]$ is the matrix exponential of $\mathbf{J}^{(0)}$. The matrix $\mathbf{J}_c(t)$ will be composed of linear combinations of the autocorrelation functions $\langle \alpha_i(t)\ \alpha_j(0) \rangle$. Replace each of these by the $(i,j)^{th}$ element of the matrix $\exp[\mathbf{J}^{(0)}\ t] \cdot \boldsymbol{\Xi}$,

$$\langle \alpha_i(t)\ \alpha_j(0) \rangle = [\ \exp[\mathbf{J}^{(0)}\ t] \cdot \boldsymbol{\Xi}\ ]_{ij}, \tag{38}$$

(see Eq. 12 above).

6. The correction matrix $\mathbf{J}_c(t)$ is composed of exponential terms of the form e^{at}, facilitating the computation of the Laplace transform $\hat{\mathbf{J}}_c(\lambda')$. Simply replace each term e^{at} with $(\lambda' - a)^{-1}$,

$$\hat{\mathbf{J}}_c(\lambda') = \mathbf{J}_c(t)|_{e^{at} \to (\lambda' - a)^{-1}}. \tag{39}$$

7. Solve the resolvent equation for λ',

$$\det[\lambda' \cdot \mathbf{I} - \mathbf{J}^{(0)} - \frac{1}{V_{cell}}\hat{\mathbf{J}}_c(\lambda')]. \tag{40}$$

The algorithm described above is easily implemented in symbolic mathematics packages. A version coded in *Mathematica* is available from the authors upon request.

[1] W. Horsthemke and R. Lefever. *Noise-induced transitions*. Springer, 1984.

[2] M. Kerszberg. Noise, delays, robustness, canalization and all that. *Current Opinion in Genetics and Development*, 14:440–445, 2004.

[3] J. Elf and M. Ehrenberg. Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Research*, 13:2475–2484, 2003.

[4] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland-Elsevier, 1992. Chapter VII.

[5] L. N. Trefethen and M. Embree. *Spectra and Pseudospectra : The Behavior of Nonnormal Matrices and Operators*. Princeton University Press, 2005.

[6] N. G. van Kampen. Itô versus Stratonovich. *Journal of Statistical Physics*, 24:175 – 187, 1981.

[7] R. C. Bourret. Stochastically perturbed fields, with applications to wave propagation in random media. *Nuovo Cimento*, 26:1–31, 1962.

[8] R. C. Bourret. Ficton theory of dynamical systems with noisy parameters. *Canadian Journal of Physics*, 43:619–639, 1965.

[9] N. G. van Kampen. Stochastic differential equations. *Physics Reports*, 24:171–228, 1976.

[10] S. I. Grossman and R. K. Miller. Nonlinear Volterra integrodifferential systems with L^1-kernels. *Journal of Differential Equations*, 13:551–566, 1973.

[11] D. McQuarrie. Stochastic approach to chemical kinetics. *Journal of Applied Probability*, 4:413–478, 1967.

[12] N. G. van Kampen. The expansion of the Master equation. *Advances in Chemical Physics*, 34:245–308, 1976.

[13] M. Scott, B. Ingalls, and M. Kaern. Estimations of intrinsic and extrinsic noise in models of nonlinear genetic networks. *Chaos*, 16:art. 026107, 2006.

[14] R. Kubo, K. Matsuo, and K. Kitahara. Fluctuation and relaxation of macrovariables. *Journal of Statistical Physics*, 9:51–96, 1973.

[15] T. B. Kepler and T. C. Elston. Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophysical Journal*, 81:3116–3136, 2001.

[16] C. W. Gardiner. *Handbook of Stochastic Methods*. Springer, 3 edition, 2004.

[17] D. T. Gillespie. Chemical Langevin equation. *Journal of Chemical Physics*, 113:297–306, 2000.

[18] H. Risken. *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer, Berlin, 1989.

[19] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics*, 6:451–464, 2005.

[20] M. Thattai and A. van Oudenaarden. Intrinsic noise in gene regulatort networks. *Proceedings of the National Academy of Science U.S.A.*, 98:8614–8619, 2001.

[21] L. Bintu, N. E. Buchler, H. G. Garcia, U. Gerland, T. Hwa, J. Kondev, and R. Phillips. Transcriptional regulation by the numbers: models. *Current Opinion in Genetics and Development*, 15:116–124, 2005.

[22] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81:2340–2361, 1977.

[23] J. M. G. Vilar, H. Y. Kueh, N. Barkai, and S. Leibler. Mechanisms of noise-resistance in genetic oscillators. *Proceedings of the National Academy of Science U.S.A.*, 99:5988–5992, 2002.

[24] R. Tomioka, H. Kimura, T. J. Kobayashi, and K. Aihara. Multivariate analysis of noise in genetic regulatory networks. *Journal of Theoretical Biology*, 229:501–21, 2004.